

10/PCT No. 18 MAY 2001

## **Verfahren zur Identifikation von Cytosin-Methylierungsmustern in genomischer DNA**

5 Die Erfindung betrifft ein Verfahren zur Identifikation von 5-Methylcytosin-Positionen in genomischer DNA.

Die genetische Information, die durch vollständige Sequenzierung genomischer DNA als Basenabfolge erhalten  
10 wird, beschreibt das Genom einer Zelle nur unvollständig. 5-Methylcytosin-Nucleobasen, die durch reversible Methylierung von DNA in der Zelle entstehen, sind ein epigenetischer Informationsträger und dienen beispielsweise zur Regulation von Promotoren. Der  
15 Methylierungszustand eines Genoms repräsentiert den gegenwärtigen Status der Genexpression, ähnlich wie ein mRNA Expressionsmuster.

5-Methylcytosin ist die häufigste kovalent modifizierte  
20 Base in der DNA eukaryotischer Zellen. Sie spielt beispielsweise eine Rolle in der Regulation der Transkription, genomischem Imprinting und in der Tumorgenese. Die Identifizierung von 5-Methylcytosin als Bestandteil genetischer Information ist daher von  
25 erheblichem Interesse. 5-Methylcytosin-Positionen können jedoch nicht durch Sequenzierung identifiziert werden, da 5-Methylcytosin das gleiche Basenpaarungsverhalten aufweist wie Cytosin. Unglücklicherweise geht bei einer PCR-Amplifikation die epigenetische Information, die die  
30 5-Methylcytosine tragen, vollständig verloren, und es existiert kein Verfahren diese Information durch einen Amplifikationsschritt zu erhalten.

Es sind mehrere Verfahren bekannt, die diese Probleme lösen. Meist wird eine chemische Reaktion oder enzymatische Behandlung der genomischen DNA durchgeführt, infolge derer sich die Cytosin- von den Methylcytosin-Nucleobasen unterscheiden lassen. Eine gängige Methode ist die Umsetzung von genomischer DNA mit Disulfit (auch als Bisulfit oder Pyrosulfit bezeichnet), die nach alkalischer Hydrolyse in zwei Schritten zu einer Umwandlung der Cytosin Basen in Uracil führt (Shapiro, R., Cohen, B., Servis, R. Nature 227, 1047 (1970). 5-Methylcytosin bleibt unter diesen Bedingungen unverändert. Die Umwandlung von C in U führt zu einer Veränderung der Basensequenz, aus der sich durch Sequenzierung nun die ursprünglichen 5-Methylcytosine ermitteln lassen (nur diese liefern noch eine Bande in der C-Spur).

Eine Übersicht über die weiteren bekannten Möglichkeiten, 5-Methylcytosine nachzuweisen, ist mitsamt der dazugehörigen Literatur dem folgenden Übersichtsartikel zu entnehmen: Rein, T., DePamphilis, M. L., Zorbas, H., Nucleic Acids Res. 26, 2255 (1998).

In der DD 293 139 A5 wird ein Verfahren zur Charakterisierung bestimmter DNA-Sequenzen beschrieben, bei dem die DNA-Moleküle, deren nichtmethylierte Erkennungsorte durch eine entsprechende Restriktionsendonuklease geschnitten werden sollen, im Reaktionsgemisch mit einer zweiten, unmethylierten DNA-Species (vor allem Oligonukleotid-Duplexe, die den Erkennungsort enthalten) inkubiert.

Die WO 97/46705 A1 offenbart ein Verfahren zum Nachweis einer methylierten, CpG enthaltenden Nukleinsäure, wobei

- man die Nukleinsäure enthaltende Probe mit einem Reagenz in Kontakt bringt, welches unmethyliertes Cytosin modifiziert, die CpG enthaltende Nukleinsäure in der Probe mittels CpG-spezifischen Oligonukleotidprimern
- 5 amplifiziert, wobei der Oligonukleotikprimer zwischen modifizierten methylierten und nichtmethylierten Nukleinsäuren unterscheidet und die methylierten Nukleinsäuren nachweist.
- 10 Ferner beschreibt US 5,824,471 A1 ein Verfahren zur Bestimmung von Abweichungen zwischen zwei Nukleinsäuresträngen, wobei eine Vielzahl von Duplexen aus den beiden Strängen oder deren Teile gebildet und diese Duplex mit einer ersten und einer zweiten
- 15 unterschiedlichen Bakteriophagen Resolvase in Kontakt gebracht wird und wobei man dann feststellt, von welcher Bakteriophagen Resolvase die Duplex gespalten wird, wobei hierdurch die Unterschiede ermittelt werden.
- 20 Nicht immer jedoch ist es erforderlich, tatsächlich die gesamte Sequenz eines Gens oder Genabschnitts zu ermitteln, wie dies bei einer Sequenzierung das Ziel ist. Dies ist insbesondere dann der Fall, wenn wenige 5-Methylcytosin-Positionen innerhalb einer längeren
- 25 Basensequenz für eine Vielzahl von unterschiedlichen Proben abzutasten sind. Die Sequenzierung liefert hier in großem Umfang redundante Information und ist zudem sehr teuer. Dies ist auch schon dann der Fall, wenn die Sequenz bereits bekannt ist und ausschließlich
- 30 Methylierungspositionen dargestellt werden sollen. Auch ist es denkbar, daß in einigen Fällen überhaupt nur die Unterschiede im Methylierungsmuster zwischen verschiedenen genomischen DNA-Proben von Interesse sind und daß auf die Ermittlung einer Vielzahl
- 35 übereinstimmender methylierter Positionen wie auch auf

die Sequenzierung verzichtet werden kann. Für die hier angeführten Fragestellungen existiert bislang kein Verfahren, das ohne Sequenzierung jeder einzelnen Probe kostengünstig die gewünschten Ergebnisse liefert.

5

Sequenzinformation muß auch deshalb immer weniger neu ermittelt werden, weil die Genomprojekte, deren Ziel die vollständige Sequenz verschiedener Organismen ist, zügig voranschreiten. Vom menschlichen Genom sind zwar derzeit  
10 erst etwa 5 % fertig sequenziert, jedoch kommen jetzt, weil andere Genomprojekte dem Ende zuneigen und dadurch Sequenzierressourcen frei werden, jedes Jahr weitere 5 % dazu. Mit der Vervollständigung der Sequenzierung des menschlichen Genoms wird bis zum Jahre 2006 gerechnet.

15

Matrix-assistierte Laser Desorptions/Ionisations Massenspektrometrie (MALDI) ist eine neue, sehr leistungsfähige Entwicklung für die Analyse von Biomolekülen (Karas, M. and Hillenkamp, F. 1988. Laser  
20 desorption ionization of proteins with molecular masses exceeding 10.000 daltons. Anal. Chem. 60: 2299-2301). Ein Analytmolekül wird in eine im UV absorbierende Matrix eingebettet. Durch einen kurzen Laserpuls wird die Matrix ins Vakuum verdampft und der Analyt so unfragmentiert in  
25 die Gasphase befördert. Eine angelegte Spannung beschleunigt die Ionen in ein feldfreies Flugrohr. Auf Grund ihrer verschiedenen Massen werden Ionen unterschiedlich stark beschleunigt. Kleinere Ionen erreichen den Detektor früher als größere. Die Flugzeit  
30 wird in die Masse der Ionen umgerechnet. Derzeit ist diese Technologie in der Lage im Massenbereich von 1.000 bis 4.000 Da Moleküle mit einer Massendifferenz von 1 Da zu unterscheiden. Durch die natürliche Verteilung von Isotopen sind die meisten Biomoleküle jedoch schon etwa 5

Da breit. Technisch ist diese massenspektrometrische Methode also vorzüglich für die Analyse von Biomolekülen geeignet. Vernünftigerweise müssen zu analysierende Produkte, die unterschieden werden sollen, mindestens 5 Da auseinander liegen. In diesem Massenbereich könnten also 600 Moleküle unterschieden werden. Im Bereich zwischen 4.000 und 100.000 Da wird zwar nicht mehr Isotopenauflösung erzielt, jedoch ist dieser Bereich auch anwendbar. Kürzlich ist der Einsatz eines infrarot (IR) Lasers gekoppelt mit der MALDI Analyse von DNA beschrieben worden (Berkenkamp, S., Kirpkar, F. and Hillenkamp, F. 1998. Infrared MALDI mass spectrometry of large nucleic acids. Science. 281: 260-262). Durch diese Kombination wurde es möglich DNA Fragmente mit einer Größe von bis zu 2.500 Basen zu detektieren.

Chemical mismatch cleavage ist eine Methode mit welcher kleine Unterschiede zwischen zwei DNA Einzelsträngen aufgezeigt werden können (Cotton, R.G.H., Rodriguez, N.R. and Campbell, R.D. 1988. Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proc. Natl. Acad. Sci. USA. 85: 4397-4401; Cotton, R.G.H. 1993. Current methods for mutation detection. Mut. Res. 285: 125-144; Saleeba, J.A. and Cotton, R.G.H. 1993. Chemical cleavage of mismatch to detect mutations. Methods in Enzymology. 217: 286-295; Smooker, P.M. and Cotton, R.G.H. 1993. The use of chemical reagents in the detection of DNA mutations. Mutations Res. 288: 65-77). Die chemische Reaktivität von C und T gegenüber Osmiumtetroxid und von C gegenüber Hydroxylamin ist erhöht, wenn diese nicht mit ihren jeweiligen komplementären Basen gepaart sind. Durch die anschließende Behandlung mit Piperidin wird der

Nukleinsäurestrang an der modifizierten Position gebrochen.

5 Eine weitere Möglichkeit nicht komplementäre Basenpaare  
in Heteroduplex DNA aufzuzeigen, besteht im Einsatz von  
Enzymen wie Muts, die an nicht komplementäre Basenpaare  
binden (Smith, J. und Modrich, P. 1996. Mutation  
detection with MthH, MthL and MthS mismatch repair  
proteins. Proc. Natl. Acad. Sci. USA 93: 4374-4379;  
10 Parsons, B.L. und Heflich, R.H. 1997. Evaluation of Muts  
as a tool for direct measurement of point mutations in  
genomic DNA. Mut. Res. 374: 277-285).

15 Derzeit fehlt ein schnelles, kostengünstiges und  
automatisierbares Verfahren zur Auffindung von  
methylierten Cytosinen in genomischer DNA. Ein solches  
Verfahren ist aber von großem Interesse, da  
unterschiedliche Methylierungsmuster auf vielfältige  
Weise zur Charakterisierung von Zelltypen und damit zur  
20 Diagnose und Klassifizierung von Krankheiten (wie  
beispielsweise Tumoren) herangezogen werden können als  
auch zum Beispiel für Studien der Zelldifferenzierung  
genutzt werden könnten.

25 Aufgabe der vorliegenden Erfindung ist daher ein  
Verfahren zur kostengünstigen und parallelisierbaren  
Auffindung von epigenetischen Informationsträgern in Form  
von 5-Methylcytosin-Basen in genomischer DNA zu schaffen.

30 Die Aufgabe wird erfindungsgemäß durch ein Verfahren zur  
Identifikation von 5-Methylcytosin-Positionen in  
genomischer DNA gelöst, wobei man folgende  
Verfahrensschritte ausführt:  
a) man behandelt die genomische DNA einer Zelle, einer

- Zelllinie, eines Gewebes oder eines Individuums chemisch derart, daß Cytosin und 5-Methylcytosin unterschiedlich reagieren und sich in der Duplex ein unterschiedliches Basenpaarungsverhalten der beiden Produkte ergibt,
- 5 b) man amplifiziert denselben Nukleinsäure Abschnitt mittels einer Polymerasereaktion,
- c) man behandelt den gleichen Nukleinsäure Abschnitt mindestens einer weiteren Zelle, Zelllinie, Gewebes oder Individuums oder einer beliebigen Referenz-DNA
- 10 entsprechend den Schritten a) und b),
- d) man bildet aus den mindestens zwei Amplifikaten der Schritte b) und c) Heteroduplexes,
- e) man führt durch eine Reaktion, die spezifisch ist für nicht komplementäre Basenpaare, eine aufzeigbare
- 15 Markierung in die Heteroduplex ein.

- Erfindungsgemäß bevorzugt ist es, daß man es zur Identifikation von Unterschieden im Cytosin-Methylierungsmuster zwischen verschiedenen Zellen,
- 20 Zelllinien, Geweben und Individuen anwendet und nur Positionen aufzeigt, in denen die Cytosin-Methylierung zwischen unterschiedlichen Zellen, Zelllinien, Geweben oder Individuen variabel ist.
- 25 Weiterhin ist bevorzugt, daß man im Schritt a) gemäß ein Disulfit (Bisulfit, Pyrosulfit) als Reagenz zur selektiven Umwandlung von Cytosin in Uracil einsetzt, wobei 5-Methylcytosin unverändert bleibt.
- 30 Es ist ferner bevorzugt, daß man genomische DNA mehrerer Individuen, Gewebe, Zelllinien oder Zellen im Schritt b) gemeinsam amplifiziert.

- Weiterhin ist bevorzugt, daß man genomische DNA mehrerer
- 35 Individuen, Gewebe, Zelllinien oder Zellen separat

amplifiziert und anschließend gemeinsam gemäß Schritt e) behandelt.

5 Erfindungsgemäß bevorzugt ist ferner, daß man durch Bildung von Heteroduplexes aus der DNA verschiedener Individuen, Gewebe, Zelllinien oder Zellen an den Positionen Basenfehlpaarungen erzeugt, an denen in der genomischen DNA ein 5-Methylcytosin lokalisiert war.

10 Bevorzugt ist auch, daß im Schritt d) durch Bildung von Heteroduplexes mit einer vollständig methylierten Referenz-DNA an den Positionen Basenfehlpaarungen auftreten, an denen sich in der genomischen DNA Cytosin befand.

15 Bevorzugt ist außerdem, daß im Schritt d) durch Bildung von Heteroduplexes mit einer vollständig demethylierten Referenz-DNA an den Positionen Basenfehlpaarungen auftreten, an denen sich in der genomischen DNA 5-Methylcytosin befand.

20 Erfindungsgemäß ist weiterhin bevorzugt, daß die Basenfehlpaarungen mittels "chemical mismatch cleavage" (chemische Veränderung an nicht komplementären Positionen) zu einer spezifischen oder hinreichend selektiven Rückgratspaltung an diesen Positionen führen.

25 Außerdem ist bevorzugt, daß man die DNA an den Basenfehlpaarungen enzymatisch spezifisch oder hinreichend selektiv spaltet.

30 Beim erfindungsgemäßen Verfahren ist auch bevorzugt, daß man im Schritt e) 1 DNA-Fragmente erhält, deren Größe einen Rückschluß auf die Spaltungspositionen und damit auf die Position der Methylcytosine und/oder die zwischen



verschiedenen Individuen, Gewebe, Zelllinien oder Zellen  
veränderlichen Methylierungspositionen erlaubt.

5 Bevorzugt ist, daß man die Analyse der Größe  
(Molekulargewichte) der DNA-Fragmente mittels  
Massenspektrometrie durchführt.

10 Besonders bevorzugt ist dabei, daß man die Fragmente  
mittels Matrix-assistierte Laser Desorptions/Ionisations  
Flugzeitmassenspektrometrie (MALDI) analysiert.

15 Besonders bevorzugt ist ferner, daß man die Fragmente  
mittels Electrospray Ionisations Massenspektrometrie  
(ESI) analysiert.

Insbesondere bevorzugt ist es, daß man die Größe der im  
Schritt e) erzeugten Fragmente der Leistungsfähigkeit des  
Massenspektrometers anpaßt.

20 Ganz besonders bevorzugt ist es, daß man mehrere PCRs  
eines Genabschnittes ausführt und die Primer schrittweise  
derart neu setzt, daß die zu erwartende Fragmentgröße  
jeweils mindestens in einer dieser PCRs in den mittels  
Massenspektrometrie nachweisbaren Massenbereich fällt.

25 Besonders bevorzugt ist es, daß man einen der PCR-Primer  
schrittweise um den maximal nachweisbaren Massenbereich  
des Massenspektrometers relativ zu dem anderen neu  
positioniert.

30 Erfindungsgemäß bevorzugt ist es, daß man im Schritt b)  
einen Primer der PCR mit einer chemischen Funktion  
versieht, so daß sich das PCR Produkt an einer Oberfläche  
immobilisieren läßt.

35

Erfindungsgemäß bevorzugt ist es auch, daß man das im Schritt b) hergestellte PCR Produkt in verschiedene Reaktionsgefäße überführt und die Oberflächen der Reaktionsgefäße chemisch so geartet sind, daß das PCR  
5 Produkt daran gebunden werden kann.

Besonders bevorzugt ist es auch, daß man im Schritt c) hergestellte PCR Produkte verschiedener Individuen in verschiedene, wie oben beschrieben zubereitete  
10 Reaktionsgefäße überführt.

Außerdem ist erfindungsgemäß bevorzugt, daß man für den Schritt e) ein Enzym einsetzt, welches mit einem nicht komplementären Basenpaar einen Komplex bildet.  
15

Ganz besonders ist dabei bevorzugt, daß dieses Enzym Muts ist.

Weiterhin ist bevorzugt, daß das Enzym eine Markierung trägt, durch welche ein Komplex veranschaulicht werden kann.  
20

Bevorzugt ist erfindungsgemäß auch, daß die Markierung eine Fluoreszenzmarkierung, eine  
25 Chemilumineszenzmarkierung, ein Massentag oder ein photochemisch abspaltbares Massentag ist.

Weiterhin ist erfindungsgemäß bevorzugt, daß man eine amplifizierte DNA-Probe gemäß Schritt c) in Anspruch 1, welche einen Unterschied zu einer amplifizierten DNA-Probe im Schritt b) in Anspruch 1 aufzeigt, in einem zweiten Durchlauf des Verfahrens selbst eine DNA-Probe nach Schritt b) in Anspruch 1 wird und mit allen anderen zu untersuchenden DNA-Proben vergleicht.  
30

35

Auch ist erfindungsgemäß bevorzugt, daß man eine Vorselektion der massenspektrometrisch im Detail zu untersuchenden Genabschnitte über eine Fluoreszenzmarkierung oder Chemiluminiszenzmarkierung des immobilisierten DNA-Stranges durchführt, deren Fehlen nach Durchführung der Schritte d) und e) des Anspruchs 1 und eines Waschschriffs das Vorhandensein von methylierten Cytosinen im untersuchten genomischen DNA-Abschnitt anzeigt.

10

Ferner ist erfindungsgemäß bevorzugt, daß man eine Vorselektion der massenspektrometrisch im Detail zu untersuchenden Genabschnitte über eine unspezifischere Variante gemäß den Ansprüchen 20 bis 23 durchführt.

15

Ein weiterer Gegenstand der vorliegenden Erfindung ist ein Kit zur Durchführung des erfindungsgemäßen Verfahrens, umfassend DNA von mindestens zwei möglichst verschiedenen Individuen, Geweben, Zelllinien oder Zellen sowie Reagenzien um die variablen Methylierungspositionen aufzuzeigen.

20

Das erfindungsgemäße umfaßt ferner vollständig methylierte und/oder demethylierte DNA und Reagenzien, die zum Nachweis von methylierten Cytosinen in einer beliebigen DNA-Probe erforderlich sind.

25

Das erfindungsgemäße Verfahren dient zur Identifikation von 5-Methylcytosin-Positionen in genomischer DNA, die unterschiedlichsten Ursprungs sein kann. Die genomische DNA wird zunächst chemisch so behandelt, daß sich ein Unterschied in der Reaktion der Cytosin-Basen zu den 5-Methylcytosin-Basen ergibt. Mögliche Reagenzien sind hier z. B. Disulfit (auch als Bisulfit oder Pyrosulfit bezeichnet), Hydrazin und Permanganat. In einer bevorzugten Variante des Verfahrens wird die genomische

30

35

DNA mit Disulfit in Gegenwart von Hydrochinon oder Hydrochinonderivaten behandelt, wobei selektiv nach anschließender alkalischer Hydrolyse die Cytosin-Basen in Uracil umgewandelt werden. 5-Methylcytosin bleibt unter diesen Bedingungen unverändert. Nach einem Aufreinigungsprozeß, der der Abtrennung des überschüssigen Disulfits dient, wird nun ein bestimmter Abschnitt der vorbehandelten genomischen DNA in einer Polymerasereaktion amplifiziert. In einer bevorzugten Variante des Verfahrens wird hier die Polymerasekettenreaktion eingesetzt. Anschließend wird derselbe Abschnitt einer anderen genomischen DNA-Probe gleichermaßen amplifiziert. Die beiden Amplifikate werden zusammengegeben, wodurch sich partiell Heteroduplexes ausbilden. In einer bevorzugten Variante des Verfahrens erfolgt dies derart, daß einer der PCR-Primer eine zur Immobilisierung geeignete Funktion trägt und daß nur ein Strang aus dem Amplifikat der ersten Probe immobilisiert wird und anschließend eine Hybridisation mit dem Amplifikat der zweiten Probe erfolgt. In einer weiteren bevorzugten Variante wird eine Vielzahl unterschiedlicher Amplifikate desselben Nukleinsäure-Abschnittes auf diese Art und Weise gegen den immobilisierten Einzelstrang aus dem Amplifikat der ersten Probe hybridisiert, welches auf viele Wells einer Mikrotiterplatte verteilt wurde. In jedem Well kann nun ein Hybridisationsexperiment durchgeführt werden.

Nach der Hybridisation wird ein Verfahren durchgeführt, das an den Positionen, in denen eine Basenfehlpaarung in der Heteroduplex auftritt, eine nachweisbare Markierung hinterläßt. In einer bevorzugten Variante des Verfahrens wird dies durch Chemical Mismatch Cleavage durchgeführt, das an den Positionen, an denen eine Basenfehlpaarung

auftritt, zu einem Rückgratbruch führt. Die dadurch erhaltenen Fragmente lassen sich durch ein beliebiges Verfahren, das die Größe von DNA-Fragmenten aufzeigen kann, analysieren. Ein solches Verfahren sollte

5 idealerweise Rückschlüsse auf jede Position in dem amplifizierten Nukleinsäureabschnitt der Probe erlauben, an denen eine Basenfehlpaarung in der Heteroduplex auftrat. Basenfehlpaarungen in der Heteroduplex sind insbesondere dann vorhanden, wenn in der DNA aus einer

10 Probe an dieser Position Cytosin vorhanden war, das in Uracil umgewandelt wurde, in der anderen jedoch 5-Methylcytosin, das bei der chemischen Vorbehandlung unverändert blieb. Das Verfahren kann sowohl für den Vergleich zweier oder mehrerer genomischer DNA Proben

15 genutzt werden, in diesem Fall liefert die Analyse der Fragmente nur die Unterschiede im Methylierungsmuster zwischen den beiden Proben im jeweiligen amplifizierten Nukleinsäure Abschnitt. Es ist aber auch möglich, eine vollständig enzymatisch am C methylierte oder

20 demethylierte DNA als Referenz einzusetzen. In diesem Fall liefert die Analyse der Fragmente alle 5-Methylcytosin-Positionen im jeweiligen amplifizierten Nukleinsäure Abschnitt.

25 In einer besonders bevorzugten Variante des Verfahren wird Massenspektrometrie für die Fragmentanalyse eingesetzt. Die Fragmente können nach vorheriger Aufreinigung im MALDI-Massenspektrometer analysiert werden. Alternativ können die Lösungen mit Electrospray

30 Ionisations Massenspektrometrie (ESI) analysiert werden. Je nach Leistungsfähigkeit der Methode und des eingesetzten Instrumentes kann es erforderlich sein, den betreffenden Nukleinsäureabschnitt in mehreren Teilschritten zu untersuchen, indem in mehreren PCRs ein

Primer schrittweise neu positioniert wird und sich damit unterschiedliche Teilamplifikate ergeben ("Primer Walking").

- 5 In einer Variante des Verfahrens können die Basenfehlpaarungen - alternativ zur Analyse von Fragmenten nach einer Rückgratspaltung in der Heteroduplex - auch durch ein Enzym nachgewiesen werden, das mit einem nicht komplementären Basenpaar einen
- 10 Komplex bildet. In einer bevorzugten Variante ist dieses Enzym MutS, das eine Markierung, z. B. eine Fluoreszenz-, Chemilumineszenz- oder Massenmarkierung, trägt.

- 15 In einer weiteren Variante des Verfahrens wird das Vorhandensein von Basenfehlpaarung, das heißt in diesem Fall auch von relevanter Information in dem amplifizierten Nukleinsäure Abschnitt, durch eine Fluoreszenz- oder Chemilumineszenzmarkierung nachgewiesen. In einer bevorzugten Variante des
- 20 Verfahrens wird ein immobilisierter DNA Strang aus dem Amplifikat der Probe 1 an dem nicht zur Immobilisierung dienenden Ende mit einer Fluoreszenzmarkierung versehen. Mit dem Amplifikat der Probe 2 werden Heteroduplexes gebildet, die einem Chemical Mismatch Cleavage
- 25 unterworfen werden. Findet eine Rückgratspaltung am immobilisierten Strang statt, so verschwindet nach einem denaturierenden Waschschrift die Fluoreszenzmarkierung, wird der Strang nicht gespalten, so bleibt die Markierung erhalten. Nur die Amplifikate, die gespalten wurden,
- 30 werden nachfolgend massenspektrometrisch näher untersucht.

### Beispiele

## Beispiel 1

## Verfahren zur Auffindung aller methylierten Cytosin-Positionen

- 5 Die zu untersuchende genomische DNA stammend aus einer Zelllinie oder möglichst nur einer Zelle wird auf zwei Reaktionsgefäße aufgeteilt und die eine Hälfte enzymatisch entweder vollständig am Cytosin methyliert oder demethyliert. Das Enzym wird thermisch inaktiviert und anschließend werden beide Teile wieder  
10 zusammengegeben und mit Disulfit und nachfolgend Alkali behandelt. Nach einer Aufreinigung wird mittels PCR amplifiziert.
- 15 Das nun durchgeführte chemical mismatch cleavage, das für C Mismatches spezifisch ist, führt zu einer Spaltung an den Positionen einer jeweiligen Heteroduplex, an denen ein ursprünglich methyliertes C vorgelegen hat, falls einer vollständigen Demethylierung der einen Hälfte der  
20 genomischen Probe durchgeführt wurde. Umgekehrt erfolgt eine Spaltung an allen ursprünglich nicht methylierten Positionen, falls eine vollständige Methylierung der einen Hälfte der genomischen Probe zuvor durchgeführt wurde. Zur Sicherheit können auch sowohl Methylierung als  
25 auch Demethylierung als Referenz durchgeführt werden, in diesem Fall müssen diese jedoch in zwei PCRs getrennt eingesetzt werden.

- 30 Variante 1: Das obige Verfahren wird so durchgeführt, das in der PCR ein Primer eingesetzt wird, der so funktionalisiert ist, daß nach der PCR eine einfache und spezifische Immobilisierung ermöglicht wird. Die Immobilisierung erfolgt an beads oder an der Oberfläche einer Mikrotiterplatte. Dies erlaubt die einfache

Abtrennung von Bestandteilen der Polymerase- und mismatch cleavage Reaktionen. Nach der chemical mismatch cleavage Reaktion wird die Duplex thermisch denaturiert und die Lösung abpipettiert. Die DNA-Fragmente werden aus dieser  
5 Lösung auf ein reversed phase Material aufgebracht und gereinigt.

Im Massenspektrometer ergeben die Fragmente eine "Leiter" von Peaks, die auf die methylierten Positionen schließen  
10 läßt. Es fallen aufgrund der symmetrischen Methylierung an CpG Positionen theoretisch immer zwei Peaks je CpG an, die jeweils vom sense und vom antisense-Strang stammen.

Variante 2: Die Reaktionen werden in Lösung durchgeführt und eine Aufreinigung nach den einzelnen  
15 Reaktionsschritten erfolgt, wenn notwendig, jeweils über ein reversed phase Material.

Variante 3: Mehrere Individuen oder Zelltypen werden  
20 parallel untersucht. Eine Referenz-DNA wird vollständig demethyliert und anschließend mit Disulfit behandelt. Sie wird nach Aufreinigung mittels PCR amplifiziert. Dabei wird wiederum ein Primer verwendet, der eine zur Immobilisierung geeignete Funktion trägt. Die Lösung wird  
25 auf die wells einer Mikrotiterplatte verteilt und immobilisiert. Dann erfolgt eine Hybridisierung gegen die PCR-Produkte aus ebenfalls mit Disulfit behandelten Proben, jeweils eine je Well (siehe auch ausführliches Beispiel mit 97 Individuen).

30

Variante 4: In dem Fall, daß das Massenspektrometer den Meßbereich nicht abdecken kann, der für die Analyse des gesamten PCR-Produktes auf Methylierungen erforderlich wäre, kann der Bereich von Interesse auch schrittweise



abgetastet werden, indem mehrere PCRs durchgeführt werden und jeweils einer der Primer um den jeweiligen Meßbereich des Massenspektrometers näher an den anderen herangesetzt wird. Damit wird beispielsweise immer nur der Bereich  
5 erfaßt, der zwischen dem zu verschiebenden Primer der jeweiligen und der nächsten PCR liegt. Das Verfahren ist mit den anderen Varianten kombinierbar.

## Beispiel 2

10 Verfahren zur Auffindung von Positionen mit variabler Cytosin-Methylierung

DNA verschiedener Individuen oder Zelllinien wird gepoolt und wie oben beschrieben eine Behandlung mit Disulfit  
15 durchgeführt. Nach alkalischer Hydrolyse der Bisulfit-Addukte und Aufreinigung der Produkt-DNA wird diese mittels PCR amplifiziert. Anschließend wird erneut aufgereinigt und das PCR-Produkt nach einigen Minuten Reannealing bei 25°C mit OsO<sub>4</sub> an den Positionen mit einem  
20 C Mismatch gespalten (chemical mismatch cleavage). Ein Mismatch C gegen A tritt immer dann auf, wenn vor der Bisulfit-Behandlung nur in einigen Individuen dort ein methyliertes Cytosin vorhanden war. Gleichsam werden bei diesem Prozeß auch etwaige SNPs (single nucleotide  
25 polymorphisms) zur Spaltung der DNA führen. Diese müssen von den aufzufindenden Methylierungspositionen unterschieden werden, was durch das oben beschriebene Verfahren zur Auffindung aller methylierten Cytosine gewährleistet ist.

30

Die Produkt DNA wird nun massenspektrometrisch wie oben beschrieben untersucht. Wenn das anfangs generierte PCR-Produkt eine größere Länge aufweist als mit der zur Verfügung stehenden Technologie seitens der

Massenspektrometrie nachzuweisen ist, so ist es möglich, daß die durch das chemical mismatch cleavage produzierten Fragmente nicht nachweisbar sind. Um dies zu umgehen, können mehrere PCRs iterativ durchgeführt werden, das  
5 heißt das ein Primer immer konstant gehalten wird, während der andere Primer immer in mehreren Schritten jeweils um die Nachweisgrenze des Massenspektrometers näher an dem anderen Primer positioniert wird (Primer walking).

10

### Beispiel 3

#### Verfahren mit 97 Individuen

Ein genomischer Abschnitt eines Individuums  
15 (Referenzindividuum) wird mit Disulfit behandelt und damit die Cytosine nach anschließender alkalischer Hydrolyse des Bisulfitadduktes in Uracile umgewandelt. Die Methylcytosine bleiben bei dieser Reaktionsfolge unangetastet. Das Produkt wird aufgereinigt und mittels  
20 PCR amplifiziert. Einer der PCR-Primer ist am 5'Ende mit einer chemischen Modifikation versehen, die zur Immobilisierung dient. Das Produkt dieser PCR wird in die 96 Kammern einer Mikrotiterplatte gegeben und die Bindung der PCR Produkte mit der Oberfläche induziert. Da nur ein  
25 Primer mit der chemischen Modifikation für die Bindung versehen ist, bindet nur ein DNA Strang an die Oberfläche. Die Platte wird gewaschen um Reagenzien der Bindungschemie und die komplementären Stränge zu beseitigen. Somit ist die Platte mit dem Referenz-DNA-  
30 Stück vorbereitet. Der gleiche genomische Abschnitt wird in jedem der 96 anderen Individuen analog mit Disulfit behandelt und anschließend amplifiziert. Für diese PCR werden je zwei normale, unmodifizierte Primer der gleichen Sequenz wie für das Referenzindividuum

verwendet. Die 96 PCR Produkte werden in die 96 Wells der vorbereiteten Platte gegeben. Durch Aufheizen und langsames Abkühlen werden die komplementären Stränge der 96 Individuen an die Referenz DNA hybridisiert (Bildung der Heteroduplex). Die 96 Individuen und Reagenzien vorheriger Reaktionen zu beseitigen. Eine  $\text{OsO}_4$ -Lösung wird in jedes der 96 Wells zugegeben, inkubiert und anschließend mit Piperidin ein Rückgratbruch in einer Heteroduplex mit einem nicht komplementären Basenpaar, von der eine Base C ist, induziert. Dies ist immer dann der Fall, wenn nur in einem Strang der Heteroduplex, das heißt in einem der Individuen, ein Methylcytosin anstelle eines Cytosins vorgelegen hat. In diesem Fall wurde nur das Cytosin des einen Individuums vor der PCR in ein Uracil überführt, wodurch sich in der Heteroduplex mit dem Gegenstrang eines anderen Individuums ein Mismatch ergibt. Der Assay ergibt also nicht direkt alle methylierten Cytosine eines genomischen Abschnittes, sondern nur diejenigen, die zwischen unterschiedlichen Individuen, Geweben, Zelllinien oder einzelnen Zellen variabel sind.

Die Heteroduplex wird durch Aufheizen aufgeschmolzen und die Lösung in ein Massenspektrometer überführt.

25

#### Beispiel 4

#### Überführen der Lösung ins Massenspektrometer

Eine gute Variante ist die Lösung nach dem Schmelzen der Heteroduplex in eine Pipettenspitze aufzunehmen, die mit einem reversed-phase Material bestückt ist. Die DNA Produkte binden daran, indem sie hydrophobe Wechselwirkungen über ihre Trialkylammonium-Gegenionen ausbilden und können so in mehreren Waschschr

Reagenzien der chemical mismatch cleavage gereinigt werden. Mit 30 % Acetonitril können die DNA Produkte wieder vom reverse-phase Material gelöst werden. Es bietet sich dabei an, die Produkte direkt auf eine  
5 vorbereitete Matrix auf einem MALDI Target zu geben. Nach dem Eintrocknen wird das Target ins Massenspektrometer eingeführt und die Produkte analysiert.

#### Beispiel 5

10 Vorselektion mittels Fluoreszenzmarkierung von für die Methylierungsdetektion relevanten Genabschnitten

Die zu untersuchende genomische DNA wird nach der wie oben beschriebenen Bisulfit Reaktion mit nachfolgender  
15 PCR Amplifikation, bei der wiederum einer der Primer eine zur nachfolgenden Immobilisierung dienliche Funktion trägt, an Beads oder einer entsprechend beschichteten Mikrotiterplatte immobilisiert. Vollständig demethylierte DNA, wie die Proben-DNA behandelt, wird als Referenz DNA  
20 verwendet und bildet mit der immobilisierten Proben-DNA eine Heteroduplex. Dann wird enzymatisch, beispielsweise mit terminal transferase, an die 3'-Enden des Produktes eine einzelne fluoreszenzmarkierte Base angehängt. Die nachfolgend durchgeführte "Chemical Mismatch Cleavage"  
25 Reaktion führt in dem Fall, das sich ein C/A Mismatch in dem Produkt befindet, zu einer Spaltung des immobilisierten Stranges, so daß nach der thermischen Dehybridisierung alle Fluoreszenzmarkierungen im nachfolgenden Waschschrift entfernt werden. War also eine  
30 Methylierung innerhalb des Amplifikates vorhanden, so tritt keine Fluoreszenz in der Mikrotiterplatte oder am bead mehr auf. Diese bleibt nur dann erhalten, wenn kein Mismatch vorliegt und damit keine 5-Methylcytosine im fraglichen Genabschnitt vorhanden sind. Bei dem Verfahren

ist zu berücksichtigen, daß z. B. SNPs ein falsch positives Signal ergeben können.

5      Sinngemäß kann dieses Verfahren auch zur einfachen, fluoreszenzbasierten Detektion von 5-Methylcytosinen in kleinen Genabschnitten, z. B. Promotoren genutzt werden. Es kann aber nur die Aussage getroffen werden, ob in der fraglichen Region Methylierungen vorliegen oder nicht, nicht aber wieviele und an welchen Positionen. Dem steht  
10      allerdings ein relativ geringer experimenteller Aufwand und eine gute Parallelisierbarkeit entgegen.

Bei den als relevant klassifizierten Genabschnitten kann dann analog einem der oben angeführten Beispiele die  
15      genaue Position den Methylcytosine durch Massenspektrometrie ermittelt werden.

#### Beispiel 6

##### Vorselektion von Heteroduplexes mit Mismatches

20

Die in einer Mikrotiterplatte immobilisierten Heteroduplexes werden zunächst mit einer Lösung von MutS, an welches ein fluoreszenter Farbstoff gebunden ist, zusammengebracht. Nur die Gefäße, in denen sich MutS an  
25      Basenfehlpaarungspositionen angelagert hat, was sich dadurch zeigt, daß nach mehreren Waschschritten immer noch die Fluoreszenz detektierbar ist, werden nachfolgend dem chemical mismatch cleavage unterworfen und im Massenspektrometer analysiert. Dadurch wird Zeit im  
30      Massenspektrometer und die Kosten für die Aufreinigung gespart, indem die Analyse von Proben ohne nachweisbare epigenetische Information vermieden wird.

## Patentansprüche

1. Verfahren zur Identifikation von 5-Methylcytosin-  
Positionen in genomischer DNA dadurch gekennzeichnet,  
daß man folgende Verfahrensschritte ausführt:
- a) man behandelt die genomische DNA einer Zelle,  
einer Zelllinie, eines Gewebes oder eines Individuums  
chemisch derart, daß Cytosin und 5-Methylcytosin  
unterschiedlich reagieren und sich in der Duplex ein  
unterschiedliches Basenpaarungsverhalten der beiden  
Produkte ergibt,
- b) man amplifiziert denselben Nukleinsäure  
Abschnitt mittels einer Polymerasereaktion,
- c) man behandelt den gleichen Nukleinsäure  
Abschnitt mindestens einer weiteren Zelle, Zelllinie,  
Gewebes oder Individuums oder einer beliebigen  
Referenz-DNA entsprechend den Schritten a) und b),
- d) man bildet aus den mindestens zwei Amplifikaten  
der Schritte b) und c) Heteroduplexes,
- e) man führt durch eine Reaktion, die spezifisch  
ist für nicht komplementäre Basenpaare, eine  
aufzeigbare Markierung in die Heteroduplex ein.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet,  
daß man es zur Identifikation von Unterschieden im  
Cytosin-Methylierungsmuster zwischen verschiedenen  
Zellen, Zelllinien, Geweben und Individuen anwendet  
und nur Positionen aufzeigt, in denen die Cytosin-  
Methylierung zwischen unterschiedlichen Zellen,  
Zelllinien, Geweben oder Individuen variabel ist.

3. Verfahren nach den Ansprüchen 1 bis 2, dadurch gekennzeichnet, daß man im Schritt a) gemäß Anspruch 1 ein Disulfit (Bisulfit, Pyrosulfit) als Reagenz zur  
5 selektiven Umwandlung von Cytosin in Uracil einsetzt, wobei 5-Methylcytosin unverändert bleibt.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß man genomische DNA mehrerer  
10 Individuen, Gewebe, Zelllinien oder Zellen im Schritt b) des Anspruchs 1 gemeinsam amplifiziert.
5. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß man genomische DNA mehrerer  
15 Individuen, Gewebe, Zelllinien oder Zellen separat amplifiziert und anschließend gemeinsam gemäß Schritt e) des Anspruchs 1 behandelt.
6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß man durch Bildung von  
20 Heteroduplexes aus der DNA verschiedener Individuen, Gewebe, Zelllinien oder Zellen an den Positionen Basenfehlpaarungen erzeugt, an denen in der genomischen DNA ein 5-Methylcytosin lokalisiert war.
7. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß im Schritt d) durch Bildung von Heteroduplexes mit einer vollständig methylierten Referenz-DNA an  
25 den Positionen Basenfehlpaarungen auftreten, an denen sich in der genomischen DNA Cytosin befand.  
30
8. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß im Schritt d) durch Bildung von Heteroduplexes mit einer vollständig demethylierten Referenz-DNA an  
35 den Positionen Basenfehlpaarungen auftreten, an denen sich in der genomischen DNA 5-Methylcytosin befand.

9. Verfahren nach einem der Ansprüche 6 bis 8, dadurch gekennzeichnet, daß die Basenfehlpaarungen mittels "chemical mismatch cleavage" (chemische Veränderung an nicht komplementären Positionen) zu einer spezifischen oder hinreichend selektiven Rückgratspaltung an diesen Positionen führen.
10. Verfahren nach einem der Ansprüche 6 bis 8, dadurch gekennzeichnet, daß man die DNA an den Basenfehlpaarungen enzymatisch spezifisch oder hinreichend selektiv spaltet.
11. Verfahren nach einem der Ansprüche 1. bis 10, dadurch gekennzeichnet, daß man im Schritt e) gemäß Anspruch 1 DNA-Fragmente erhält, deren Größe einen Rückschluß auf die Spaltungspositionen und damit auf die Position der Methylcytosine und/oder die zwischen verschiedenen Individuen, Gewebe, Zelllinien oder Zellen veränderlichen Methylierungspositionen erlaubt.
12. Verfahren nach Anspruch 11, dadurch gekennzeichnet, daß man die Analyse der Größe (Molekulargewichte) der DNA-Fragmente mittels Massenspektrometrie durchführt.
13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß man die Fragmente mittels Matrix-assistierte Laser Desorptions/Ionisations Flugzeitmassenspektrometrie (MALDI) analysiert.
14. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß man die Fragmente mittels Electrospray Ionisations Massenspektrometrie (ESI) analysiert.



15. Verfahren nach den Ansprüchen 13 oder 14, dadurch gekennzeichnet, daß man die Größe der im Schritt e) gemäß Anspruch 1 erzeugten Fragmente der Leistungsfähigkeit des Massenspektrometers anpaßt.
- 5
16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß man mehrere PCRs eines Genabschnittes ausführt und die Primer schrittweise derart neu setzt, daß die zu erwartende Fragmentgröße jeweils mindestens in
- 10
- einer dieser PCRs in den mittels Massenspektrometrie nachweisbaren Massenbereich fällt.
17. Verfahren nach Anspruch 16, dadurch gekennzeichnet, daß man einen der PCR-Primer schrittweise um den
- 15
- maximal nachweisbaren Massenbereich des Massenspektrometers relativ zu dem anderen neu positioniert.
18. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß man im Schritt b) einen Primer der PCR mit einer chemischen Funktion versieht, so daß sich das PCR Produkt an einer Oberfläche immobilisieren läßt.
- 20
19. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß man das im Schritt b) hergestellte PCR Produkt in verschiedene Reaktionsgefäße überführt und die Oberflächen der Reaktionsgefäße chemisch so geartet sind, daß das PCR
- 25
- Produkt daran gebunden werden kann.
- 30
20. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß man im Schritt c) hergestellte PCR Produkte verschiedener Individuen in verschiedene nach Anspruch 19 zubereitete Reaktionsgefäße
- 35
- überführt.

21. Verfahren nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß man für den Schritt e) ein Enzym einsetzt, welches mit einem nicht komplementären Basenpaar einen Komplex bildet.
22. Verfahren nach Anspruch 21, dadurch gekennzeichnet, daß dieses Enzym MutS ist.
23. Verfahren nach Anspruch 21, dadurch gekennzeichnet, daß das Enzym eine Markierung trägt, durch welche ein Komplex veranschaulicht werden kann.
24. Verfahren nach Anspruch 21, dadurch gekennzeichnet, daß die Markierung eine Fluoreszenzmarkierung, eine Chemilumineszenzmarkierung, ein Massentag oder ein photochemisch abspaltbares Massentag ist.
25. Verfahren nach einem der Ansprüche 1 bis 24, dadurch gekennzeichnet, daß man eine amplifizierte DNA-Probe gemäß Schritt c) in Anspruch 1, welche einen Unterschied zu einer amplifizierten DNA-Probe im Schritt b) in Anspruch 1 aufzeigt, in einem zweiten Durchlauf des Verfahrens selbst eine DNA-Probe nach Schritt b) in Anspruch 1 wird und mit allen anderen zu untersuchenden DNA-Proben vergleicht.
26. Verfahren nach einem der Ansprüche 1 bis 25, dadurch gekennzeichnet, daß man eine Vorselektion der massenspektrometrisch im Detail zu untersuchenden Genabschnitte über eine Fluoreszenzmarkierung oder Chemilumineszenzmarkierung des immobilisierten DNA-Stranges durchführt, deren Fehlen nach Durchführung der Schritte d) und e) des Anspruchs 1 und eines Waschschritts das Vorhandensein von methylierten

Cytosinen im untersuchten genomischen DNA-Abschnitt anzeigt.

- 5 27. Verfahren nach einem der Ansprüche 1 bis 25, dadurch gekennzeichnet, daß man eine Vorselektion der massenspektrometrisch im Detail zu untersuchenden Genabschnitte über eine unspezifischere Variante gemäß den Ansprüchen 20 bis 23 durchführt.
- 10 28. Kit zur Durchführung eines Verfahrens nach Anspruch 1, umfassend DNA von mindestens zwei möglichst verschiedenen Individuen, Geweben, Zelllinien oder Zellen sowie Reagenzien um die variablen Methylierungspositionen aufzuzeigen.
- 15 29. Kit zur Durchführung des Verfahrens nach Anspruch 1 umfassend vollständig methylierte und/oder demethylierte DNA und Reagenzien, die zum Nachweis von methylierten Cytosinen in einer beliebigen DNA-
- 20 Probe erforderlich sind.

## Zusammenfassung

5 Beschrieben ist ein Verfahren zur Identifikation von 5-Methylcytosin-Positionen in genomischer DNA, welches dadurch gekennzeichnet ist, daß folgende Verfahrensschritte ausgeführt werden:

- 10 a) die genomische DNA einer Zelle, einer Zelllinie, eines Gewebes oder eines Individuums wird chemisch so behandelt, daß Cytosin und 5-Methylcytosin unterschiedlich reagieren und sich in der Duplex ein unterschiedliches Basenpaarungsverhalten der beiden Produkte ergibt,
- 15 b) derselbe Nukleinsäure Abschnitt mittels einer Polymerasereaktion amplifiziert wird,
- 20 c) der gleiche Nukleinsäure Abschnitt mindestens einer weiteren Zelle, Zelllinie, Gewebes oder Individuums oder einer beliebigen Referenz-DNA entsprechend den Punkten a) und b) behandelt wird,
- 25 d) aus den mindestens zwei Amplifikaten der Punkte b) und c) Heteroduplexes gebildet werden,
- e) durch eine Reaktion, die spezifisch ist für nicht komplementäre Basenpaare, eine aufzeigbare Markierung in die Heteroduplex eingeführt wird.